DOCUMENT RESUME

ED 125 299 FL 007 822

AUTHOR Leech, Geoffrey N.
TITLE Being Precise about Lexical Vagueness. York Papers in Linguistics., No. 6.
INSTITUTION York Univ. (England). Dept. of Language.
PUB DATE Mar 76
NOTE 18p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Definitions; Dictionaries; *Evaluation Methods; Linguistic Patterns; *Linguistic Theory; *Semantics; Testing; Vocabulary
IDENTIFIERS Labov (William); *Lexical Semantics

ABSTRACT

This paper accepts Labov's (1973) criticisms of the categorial approach, i.e., the view that linguistic units are categories which are discrete, invariant, qualitatively distinct, conjunctively defined, and composed of atomic primes, and follows Labov in attempting to develop a non-categorial (or fuzzy-categorial) approach to lexical semantics, based on experimental evidence. At the same time, it is argued that Labov's method of investigation, whereby informants name pictured objects, suffers from built-in limitations and distortions. An alternative method is proposed, in which informants respond to linguistic rather than pictorial stimuli. This method, though perhaps not superior to Labov's, at least has complementary advantages and disadvantages, so that its use, when combined with Labov's method, may be said to widen the empirical basis for the study of lexical semantics. (Author/DB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED125299

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

BEING PRECISE ABOUT LEXICAL VAGUENESS

Geoffrey N. Leech

(University of Lancaster)

1. *The Fuzzy-Categorial View of Lexical Meaning*

Labov in his article 'The boundaries of words and their meanings'[1] regrets that up to recently linguistics has been dominated by the 'categorial view', i.e. the view that linguistic units are categories which are discrete, invariant, qualitatively distinct, conjunctively defined, and composed of atomic primes. Orthodox componential analysis, as applied to meanings, is a good example of this strict categorial approach; in Labov's formula, adapted from Weinreich, a typical componential account specifies the meaning of a word in the following terms (1973:348):

(I) L(x) if C_1 and C_2 and C_3 .... and C_n
(where L is 'a sign', x is an object, C is a criterion of reference, and L(x) is read as 'L refers to x')

In this paper I shall accept Labov's criticisms of the categorial approach, and shall follow him in attempting to develop a non-categorial (or rather, fuzzy-categorial) approach to lexical semantics, based on experimental evidence. At the same time, I shall argue that Labov's method of investigation, whereby informants name pictured objects, suffers from built-in limitations and distortions. Here I propose and implement an alternative method (see appendix I) in which informants respond to linguistic, rather than pictorial stimuli. This method is probably not superior to Labov's, but at least has complementary advantages and disadvantages, so that its use, when combined with Labov's method, may be said to widen the empirical basis for the study of lexical semantics.

FL007822

Labov proposes a relaxation of the requisites of the strict categorial approach to meaning. He reports the results of informant tests carried out on the meaning of the word *cup*, and from this derives his own definition of *cup*, 'designed to register the categorial character of our lexicon' (1973:367), and yet to make allowance for semantic vagueness:

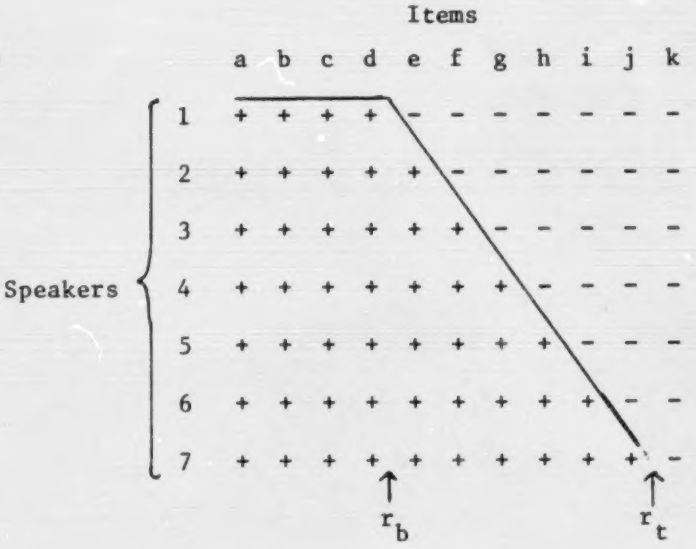
(II) The term *cup* is regularly used to denote round containers with a ratio of width to depth of $1 \pm r$ where $r \leq r_b$, and $r_b = \alpha_1 + \alpha_2 + \ldots \alpha_v$ and α_i is a positive quality when the feature i is present and 0 otherwise.

Feature 1 = with one handle
2 = made of opaque vitreous material
3 = used for the consumption of food
4 = used for the consumption of liquid food
5 = used for the consumption of hot liquid food
6 = with a saucer
7 = tapering
8 = circular in cross-section

Cup is used variably to denote such containers with ratios of width to depth of $1 \pm r$ where $r_b \leq r \leq r_t$ with a probability or $r_t - r / r_t - r_b$. The quantity of $1 \pm r_b$ expresses the distance from the modal value of width to height. (1973:366-7)

This definition is best understood by reference to an idealised and simplified diagram which Labov provides (1973:368):

(III)



The diagram shows a set of items (which we may think of as cups or cuplike objects) ordered in terms of their deviation from a width-depth ratio of 1. A plus indicates that a given object is called a 'cup' by a given speaker. Moving from the 'ideal' width-depth ratio of 1 on the left, a certain deviation (up to r_b) is tolerated without any dissent from the designation of the object as a *cup*. Beyond that point, and up to the degree of deviation r_t, however, there is progressive avoidance of the designation *cup*, and beyond this point, the term *cup* is not used. In this definition, therefore, the set of objects called cups constitutes a fuzzy set, some objects being more markedly members of the set of cups than others.

Without going into all the details of the definition, we may note that it is 'categorial' to the extent that a number of discrete criteria for reference are used. But of these, only 'roundness' and 'containerhood' are absolute and invariant. The other criteria are variant, their semantic function being specified by numerical values. Primacy in the determination of 'cuphood' is given to the continuously variable criterion of width-depth ratio. The definition specifies the referential range of *cup* as an 'invariant core' (of items which will always be called cups) and a 'variable skirt' (of items which will only sometimes be called cups).

The difference between the Labov position and the strict categorial position is that Labov gives an account of semantics in terms of *degrees*

of membership of a category. Max Black (quoted by Labov, p.353) says that a given term is vague if there exist objects concerning which it is impossible to say whether the word does or does not denote. In Labov's model, some cups are more cuplike than others. If a cup is less than 100% cuplike, we can at the most indicate the probability of denotation in a particular instance.

2. *An Evaluation of Labov's Approach*

Labov takes lexical semantics in a direction in which I like to see it going. Many semanticists have long acknowledged lexical vagueness as intuitively selfevident in a reluctant sort of way, but have failed to give it a systematic place in their descriptions, simply because it is not amenable to treatment in strict categorial/structural terms. Another school of thought (represented philosophically by Wittgenstein and Quine[2], and by linguists of a 'meaning = use' persuasion) has taught us to view any systematicisation of lexical meaning with scepticism. Strictly maintained, the view that lexical meaning is unsystematic and non-rule-governed must lead to the conclusion that meaning is arbitrary not in the Saussurean sense, but in the more radical Humpty-Dumptian sense: 'When *I* use a word, it means just what I choose it to mean, neither more nor less'. The in-between position of Labov strikes me as intuitively plausible, and can be justified by arguments based on common observation:

Arguments against the strict-categorial position:

(a) Speakers of a language often use ad hoc or strategic definitions, which essentially sharpen the boundaries, for a given communicative purpose, or an otherwise fuzzy meaning.[3]

(b) Serious disagreements about the meaning of words occur: disagreements which cannot be easily explained in terms of idiolectal variation. It seems that in such cases, two speakers present alternative 'rationalisations' (or defuzzifications) of a fuzzy definition.[4]

(c) Historical semantic change is frequently explicable in terms of changes in the criteriality of semantic features.[5]

Against the sceptical, non-systematic view of lexical meaning:

The only argument one can use here is that human beings *do* communicate more or less efficiently (e.g. that I do share with other speakers of English the conviction that it is appropriate to use the word *cup* to refer to a cup, rather than to refer to a toothpick).

I shall therefore take for granted, from now on, that Labov's approach to lexical definition is essentially correct. However, without detracting from the importance of Labov's study, I maintain that it has certain weaknesses:

(a) It fails to differentiate between referential (denotative) vagueness and semantic (connotative) vagueness.

(b) Its research method applies only to words with physical reference and to physical criteria of reference.

(c) It restricts informant testing to the procedure of naming objects.

To consider the first two of these further:

(a) *Confusion of Referential and Semantic Vagueness*
Referential and semantic vagueness may be distinguished as follows: Vagueness in Black's sense (undecidability of denotation) can have two sources. Referential vagueness arises preeminently with scalar terms such as 'old' and 'young'. The word *old*, as applied to persons, designates a fuzzy set, because people who can be designated *old* are not old to the same degree. A man of 80 will be more decidedly a member of the set of old people than a man of 60. For maximum generality, referential vagueness is best formulated not as a relation between a fuzzy set and a word, but as a relation between a fuzzy set of 'objects' (real-word phenomena) and a semantic feature, denotatively interpretable as a criterion of reference. To see this, we may note that the feature 'oldness' is present not only in the meaning of the adjective *old* itself (which is defined exclusively in terms of this feature), but is also a component of meanings of other words (e.g. *sage, veteran, dotard*). The referential vagueness of 'old' also applies to these words (e.g. all other things being equal, a person of 80 is more readily called a *dotard* than a person of 60). But one can give an account of such referential vagueness, and still preserve the format for categorial definition given in (I). For example, one could give a first-approximation definition of *sage* as follows:

Sage (x) if C_1 and C_2 and C_3

(where C_1 = 'person', C_2 = 'wise', C_3 = 'old')

Here, although the feature C_3 defines a fuzzy set, the definition adheres to the categorial view of meaning in so far as the features C_1, C_2, C_3 are invariant.

But on another level, the definition itself is a fuzzy set: a fuzzy set of semantic features. I assume, in accord with some aspects of Labov's investigation, that one component of a definition may differ in the degree of its criteriality from another; i.e. may have a greater or lesser degree of membership of the feature-set which comprises the definition. It is reasonable to surmise that some features are more central to the meaning of *sage* than others: e.g. that 'oldness' is less criterial than 'wisdom'. Variable criteriality in this sense may be termed SEMANTIC VAGUENESS.

It appears that Labov has constructed his tests in such a way as to obscure the difference between these two types of vagueness. He presented his informants with various pictures of cups or cup-like objects. Some of these were discriminated in terms of discrete properties (such as having or not having a handle), but the most important discriminatory parameter was the continuously variable one of width-depth ratio. In his definition of *cup* (II), where he assigns (unspecified) numerical values to features 1-8, it is difficult to decide whether he intends these values to represent variable criteriality (semantic vagueness) or variable reference (referential vagueness), although on the whole the features do not lend themselves to the latter interpretation.[6] On the other hand, it is clear that the variability of width-depth ratio falls into the category of referential

vagueness: in terms of this feature, different objects are assigned different degrees of 'cuphood', just as in terms of fuzzy-set theory, the feature of 'old age' would assign to people of varying ages different degrees of membership of the category 'sage'. It appears that Labov has not considered the criteriality of width-depth ratio in relation to other features; he seems to have assumed, on a priori grounds, the primacy of this criterion over others, even though other features in his list could have been similarly treated in terms of continuous scales (e.g. taperedness, roundness). The results he presents therefore do not discriminate between two variables: (a) the extent to which the criterion of width-depth is criterial to cuphood, and (b) the extent to which a given object conforms to that criterion.

(b) *Restriction to physical objects and physical properties*
Labov's technique of presenting his informants with objects (or rather pictures of objects) and asking them to name them obviously restricts his research method to physical, visually presentable objects. This restriction is unfortunately a common one: semantics has been too often limited in the past by the assumption that referring to an object somehow represents the essential paradigm for meaning in general. Although there is no reason why the technique should not be extended to non-visual sensory stimuli, there appears to be no way in which it could be applied to the investigation of abstract meanings.

A related, but less obvious, drawback of this research method is that it excludes, or at least disfavours, non-physical criteria of reference. Labov partially circumvents this difficulty by getting his informants to imagine the pictured object containing coffee, mashed potatoes, flowers etc, thereby taking account of functional, as well as physical criteria. However, the result of this stratagem is probably to mispresent functional criteria as less criterial than they really are. A frequent misunderstanding of the nature of functional criteria is to suppose that they define what an object is used for, rather than what it is *meant* to be used for. In emergencies, I may use a cup as an ash-tray or as a chamber-pot, but it doesn't stop being a cup for all that. It is arguable that for words denoting artefacts, functional features (understood in purposive terms) are in some sense more 'basic' than physical features; e.g. that in the definition of *cup*, the fact that a cup has a particular functional role in society as a drinking-vessel is more important than its having a particular width-depth ratio. I argue this from two kinds of evidence from common experience: (a) the evidence of words (like *telescope, razor*) denoting objects which are physically dissimilar, but have a common function, shows that for at least some words denoting physical objects, a reliance on physical criteria would lead to the necessity of a disjunctive definition. (b) the evidence of 'teleological *why*-questions' (as asked by young children) suggests that functional features often have an explanatory, 'logical' priority over physical features. For example, imagine an exasperating six-year-old asking the questions:

i Why is a cup hollow in the middle?
ii Why does a cup have a handle?
iii Why does a cup have a saucer to stand on?
iv Why does a cup taper towards the bottom?
v Why doesn't a cup have a spout?
vi Why is a cup for drinking out of?

In answer to questions (i) - (v), a long-suffering grown-up could give, and probably would give, a 'folk-teleological' answer, directly or indirectly

in terms of a cup's function as a drinking vessel, especially for hot liquids. But question (vi) (the most exasperating of all) is unanswerable, except metalinguistically, e.g. in terms of the etymology of the word *cup*. Evidently this is because (vi) is itself a question about a functional feature, and cannot be explained in terms of any logically prior notion of what *cup* means. Contrast:

A: Why does a cup have a handle?
B: Because it's used for drinking out of.

A: Why is a cup used for drinking out of?
B: ?* Because it has a handle.

3. *An Alternative Research Strategy: Linguistic-Stimulus Tests*

Enough has been said to indicate my reservations about Labov's method. I shall now consider the advantages and disadvantages of an alternative research strategy, in which informants respond to linguistic, rather than visual stimuli. Informant tests using linguistic stimuli are far from new (acceptability tests, for example, fall into this general category). The most direct antecedents to the present proposal are the semantic tests reported in Leech 1969, and Leech and Pepicello 1972, which were, however, undertaken within a strict-categorial framework. After describing and evaluating the method, I shall outline the theoretical framework it presupposes.

In this method, the informant is required to give a truth value rating to a statement. In one variant of the test, an IMPLICATION AND INCONSISTENCY test, (see Leech 1969), subjects are presented with two statements, *X* and *Y*, and are asked to judge the truth of *Y*, assuming that *X* is true; e.g.

X: The object that Susan is holding is a cup.
Y: The object that Susan is holding has a handle.

In another variant, a TAUTOLOGY AND CONTRADICTION test, subjects are simply presented with a statement, and asked to judge its truth value; e.g.

Cups have handles.

Previous exploratory experiments have suggested that a three-point scale

A Has to be true
B Can be either true or false
C Can't be true

can be quite sensitive in indicating degrees of criteriality, since the number of people answering 'Has to be true' appears to decline significantly as the statement becomes (in an intuitive sense) less analytic.[7] However, if we take a fuzzy-set approach to semantic features, a more delicate rating-scale e.g. a 7-point scale explicitly registering degrees of truth and falsehood,[8] may be considered more appropriate.

Among the advantages of tests using linguistic stimuli is that unlike Labov's tests, they can be used in the investigation of abstract criteria; e.g.:

{ X: The person who wrote this book is a sage.
Y: The person who wrote this book is wise.

A second advantage is that such tests can give the same treatment to physical and functional criteria; e.g.:

{ X: Susan is holding a cup.
Y: Susan is holding a drinking vessel.

A further point is that referential and semantic vagueness are separated in this test, since instead of requiring the informant to apply a word to a referent, we require him to judge the relation between one linguistic expression and another, thereby placing the emphasis on *semantic* vagueness. Let us suppose we want to test the criteriality of the component 'well-known' in the meaning of *notorious*. We can test the implication relation between *X* and *Y*, where *X* is *Fred is notorious* and *Y* is *Fred is well-known*, without bringing into play the referential fuzziness of the feature 'well-known' itself. If the two words *notorious* and *well-known* share this feature, the vagueness of 'well-known' will presumably remain constant in the two statements tested.

A final point in favour of this proposal is that the testing technique can be extended to word-classes other than nouns. For example, informants' assignments of truth-value to the statement

Things that fly are heavier than air

would provide evidence for the criteriality of one feature in the meaning of the verb *fly*.

4. *Contextual Variability*

A major stumbling-block to the systematic investigation of lexical semantics is the variability of meaning according to linguistic or extra-linguistic context. But it is important, here as elsewhere, to distinguish between referential and semantic vagueness. It is easy to cite extreme cases of referential variability according to context: e.g. a married couple may move from a housing-estate where they are considered 'old' 'well-off', to a different area of the city, where their neighbours consider them 'young' and 'poor'. This type of variability is fortunately not at issue in the linguistic-stimulus tests proposed here. But semantic vagueness involves its own kind of contextual variation. When one accepts the notion of lexical definition as a fuzzy set of semantic features, one must also accept that the criteriality of a feature may vary from one context to another. This variation can be tested, so long as we restrict attention to linguistic rather than extralinguistic context. We can, for example, investigate the effect of selection restrictions on the criteriality of features in the meaning of the adjective *veteran*. Suppose that features such as 'old' and 'of long service' are constant in such contexts as *veteran guardsman*,

veteran boxer, *veteran car*, but that the feature 'experienced', which is present in the first two collocations, is absent in the third. Such shifts in criteriality, including the contextually-activated 'suppression' of features, can be experimentally investigated by the appropriate wording of statements to be tested. It is arguable, moreover, that any variations dependent on extra-linguistic context (excluding idiolectal variations) can be matched by equivalent variations dependent on linguistic context, and hence can be studied by this method. Hence linguistic-stimulus tests can be constructed so as to be sensitive to contextual variation.

5. *Disadvantages of the Linguistic-Stimulus Method*

Some disadvantages of the linguistic-stimulus tests will be mentioned later, in discussing the results of a preliminary test listed in Table IV below. However, two obvious problems may be mentioned in advance. First, this method assumes that a lexical definition can be represented in the form of a set of features of varying criteriality, and that it is possible, by providing appropriate statements as stimuli, to separate out the features, and to assign degrees of criteriality to each of them. If the resulting definition is to be complete, this requires a further assumption that every feature of meaning can be represented in overt verbal form, i.e. that for every feature C there is at least one linguistic expression E such that the meaning of E consists solely of the feature C (with 100% criteriality). (This assumption is also made by Labov, to the extent that he represents his definition of *cup* (II) in natural language.) Although this is a very large assumption to make, I believe it is substantially correct (see the theoretical discussion in 8 below). To the extent that it is incorrect, the results of tests are incomplete (i.e. the meaning of a word is not fully decomposed into 'atomic' features) rather than incorrect.

A second problem is raised by polysemy and (more acutely) by homonymy. If a word *W1* whose meaning is being investigated is homonymous with another word *W2*, there is no guarantee that the informant's response to *W1* is not influenced by *W2*. Thus the method cannot be satisfactorily applied to homonymous words, unless the informants' attention is restricted by explicit verbal directions to the word in question, a gambit which itself runs the risk of prejudicing meaning. With polysemy, the same sort of problem arises, but is less acute, because the fuzzy-set model itself encourages one to revise one's notion of polysemy, by reinterpreting polysemy as contextual variability. Thus in the discussion of *veteran* (adj.) above, *veteran car* and *veteran boxer* were treated as instances of the same sense of *veteran*, whereas in a strict-categorial model, they would have to be assigned to different definitions, one being treated as a 'metaphorical extension' of the other.

6. *A Preliminary 'Platitude Test'*

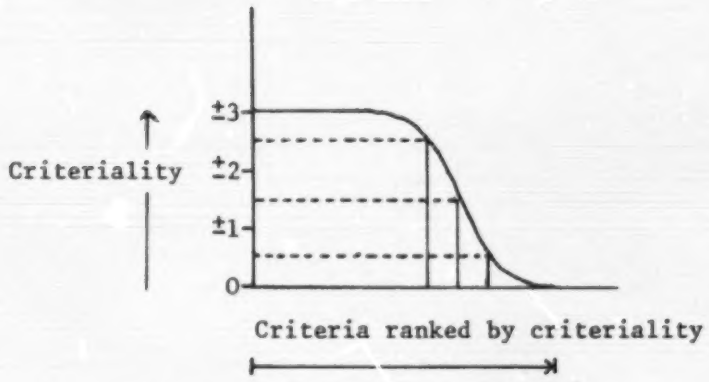
Before discussing these and other problems further, I now present results of a test which was carried out with 60 first-year undergraduates at the University of Lancaster, and which was a first attempt at the application of the proposed method.[9] The object of the test was to arrive at a tentative

fuzzy-set definition of *cup*, by asking informants to judge the truth-value of thirty generic statements about cups. The test was called a 'platitude test', since many of the stimulus sentences were essentially platitudes about cups, (or statements of a high criteriality for cuphood) such as *Cups have handles*. Such platitudes, however, were mixed with statements of low or non-existent diagnostic value for cuphood (such as *Cups are white*) and statements of negative diagnostic value, such as *Cups are musical instruments*. Statements with overt quantifiers (such as *All cups have handles*) were avoided, as it was thought that within a fuzzy-set conception of lexical meaning, generic quantification (the fuzzy equivalent of universal quantification) was exactly what was needed. By using a seven-point scale, it was hoped to register people's reactions with a greater delicacy than would be obtained (for example) by the three-point scale employed in previous work. To simplify the task for the informant, it was decided to give each point on the scale a verbal label. But inevitably, the actual choice of verbal labels was problematic. The idea of using labels for degrees of truth-value ('absolutely true'/'very true', etc.) was rejected because of informants' difficulties in interpreting such labels. Instead, labels referring to frequency were preferred:

always true +3	always false -3
generally true +2	generally false -2
often true +1	often false -1
sometimes true and sometimes false 0	

Responses were scored on the range +3 to -3 according to which label was preferred. Two doubtful features of this rating-scale are to be noted: (a) Whereas the fuzzy-set model of lexical definition strictly requires the elicitation of fuzzy truth-values such as *very true*, the test actually elicited a subjective judgement of the probability of the truth/falsehood of a given statement, in terms of a presumed two-valued logic. This discrepancy between the theoretical object of the test and its practical implementation may not be serious: a lexical definition in terms of a fuzzy-set model may, in fact, be isomorphic with one in terms of a probabilistic model. (b) There is no guarantee that the labels define equal intervals on the scale ranging from 'always true' to 'always false'. In fact, the neutral label 'sometimes true and sometimes false' logically spreads over the whole intermediate area from +2 to -2. Although subjects were required to choose only one label (that judged to be most appropriate) for each statement, the broad semantic range of the neutral label perhaps accounts for one inequality observed in the overall pattern of responses, viz. the tendency to avoid the labels 'often true' and 'often false'. On the other hand, this avoidance is compatible with, and indeed predictable from, an equal-interval scale, if one takes the Zadeh function for the membership of a fuzzy set as applicable to lexical definition (i.e. as applicable to semantic vagueness as well as to lexical vagueness):[10]

(IV)



The horizontal dotted lines on the diagram show the mid-points between the scores ±3, ±2, ±1, and 0, and thus delimit, in theory, the areas of criteriality eliciting those respective scores as responses. As the diagram indicates, the band of criteria eliciting ±1 scores would theoretically be relatively small.[11]

These problems associated with scaling and scoring obviously point to the need for the investigation and comparison of alternative scaling methods, but are unlikely to affect the broad results of the test, used simply as a means of ranking features in terms of their criteriality. An equal-interval scale would be more crucial if one were interested in measuring dispersion of responses.

The table below ranks the thirty statements tested in order of observed criteriality, assigning to each an index of criteriality within the range +1 to -1. The index is calculated simply by adding the scores (giving maximum score of 180 for 60 informants), and dividing the aggregate by 180.

(V) *Ranked list of criteria with indices of criteriality*

1	A cup is hollow in the middle	.93
	A cup is a physical object	.93
3	A cup has an open top	.83
4	A cup is a drinking vessel	.82
5	A cup is a container (i.e. an object for containing things)	.80
6	A cup is for drinking out of	.77
7	A cup is a vessel (i.e. an object for containing liquid)	.75
8	A cup has a flattish bottom	.74
9	A cup is a man-made object	.71
10	A cup has a handle	.67
11	A cup is round seen from the top	.64
12	A cup gets slightly narrower towards the bottom	.43
13	A cup is made of china or some other kind of pottery	.39
14	A cup is breakable	.37
15	A cup has a saucer to go with it	.36
16	A cup is a container (or vessel) of middling depth	.17
17	A cup is a deep container (or vessel)	.15
18	A cup is for drinking hot liquid out of	.11
	A cup is made of china	.11
20	A cup is white	-.03
21	A cup is a shallow container (or vessel)	-.18

22	A cup is blue	-.23
23	A cup is made of paper	-.33
24	A cup is made of glass	-.38
25	A cup leaks	-.47
26	A cup is for putting flowers in	-.61
27	A cup is for eating out of	-.62
	A cup has two handles	-.62
29	A cup is a musical instrument	-.84
30	A cup has an open bottom	-.93

7. Discussion of the Form and Results of the Test

These thirty statements were chosen to include those features which were judged to be most criterial in the meaning of *cup*, as well as a number of features of insignificant criteriality to act as 'distractors'. Since the number of statements which might be made about cups is infinite, this cannot count as more than a partial investigation of the word's meaning, though hopefully it includes all the more significant features. The statements are given here in the form in which they were presented to informants (except that in the test they were presented in random order). The wording of some statements was deliberately circumlocutionary, to ensure maximum clarity. Clarity also recommended the avoidance of negative statements (such as *A cup doesn't leak*), but all negative criteria can be trivially converted into positive criteria expressed in negative form. (Negative criteriality is, of course, just as important to a word's meaning as positive criteriality.) The term 'clarity' begs the question of what meaning was intended to be expressed by these statements. I assume that any feature of meaning can be represented in the form of a generic statement,[12] which, in a true formalization of lexical definition, would be expressed in terms of a semantic/logical formula. For example, a rough formal representation of statement 6 *A cup is for drinking out of* might be given as follows (where ⌝ represents the generic [as distinct from universal] operator):

$$\neg x\ ((\text{Cup } x) \supset (\exists y)\ (\text{Man } y \ \&\ \text{Made } y,x \ \&\ \text{Intend } y,\pi)))$$
$$\text{Where } \pi = (\exists z)\ (\text{Drink-out-of } z,x)$$

Regrettably, the English-language wording of statements does not always precisely capture my intended logical content, as was made clear by one or two informants' comments after the test. Some difficulty was felt in the interpretation of functional statements such as 6 *A cup is for drinking out of*, as the following comment by an informant shows: 'The phrase "A cup is for..." is ambiguous -- because it can mean made for, apt for or can be used for... So the response must tend to be erratic.' A probable consequence of this lack of clarity is an underweighting of functional criteria in the test results. Even so, functional criteria (e.g. 4, 5, 6, 7) were amongst those with a high index of criteriality, and to that extent my suspicion that Labov's tests were biased against such criteria received support.

Another difficulty was noted in the use of general superordinate terms such as *vessel* and *container*. Although it was intended that statements like 5 *A cup is a container* and 7 *A cup is a vessel* would be generally interpreted as tautologies, their indices of .10 and .75 respectively

showed that this was not always the case. A probable reason for this is that suggested by the informant who commented: 'When you first said that a cup was a container, I said this was false since I felt a container to be something which holds something in it over a long period of time'. Similarly, a 'vessel' may be understood as something in which one *keeps* liquid, rather than something in which one holds it for the purpose of drinking. Presumably, this feature of vesselhood is suppressed in the collocation *drinking vessel*, with the result that statement 4 elicited a more categorical response than statement 7, which it appears to entail. These observations tend to suggest (as, indeed, is not surprising if one abandons a strict-categorial approach) that true hyponyms are rare in natural language, and that the core meaning of the superordinate term does not necessarily include that of its hyponym.

One general conclusion from the test is that informants apparently find responding to specific, physical criteria an easier and more straightforward task than reacting to more abstract criteria, where the vagueness factor seems to be more pervasive. As a consequence, the intuitive logical priority of one feature over another in terms of entailment or 'folk-teleology' (see 2 above) is not always reflected in the ranking of the test-statements (e.g. compare criteria 4 and 1 in (V)). Moreover, because of these and other interference factors (e.g. inattention or error by informants) the maximum criteriality of 1.00, which I assume would not be of infrequent occurrence in an 'ideal' test, will rarely if ever appear in an actual test.

It is evident, then, that the linguistic-stimulus test (at least as illustrated in this exploratory version), produces an 'observational criteriality' which falls short of the 'ideal criteriality' of a complete and precise representation of what a word like *cup* means to a body of native speakers. Nevertheless, this may reasonably count as a first approximation to an empirically-arrived-at definition of *cup*, according in the main with one's intuitive assessment of the relative importance of features to the concept of cuphood. How good the approximation is must be left, at this stage, to the reader's judgement.

8. *Theoretical Considerations*

The claim that a semantic test provides an approximation to the meaning of an English word begs a number of questions about semantic theory, and the theoretical justification for testing.

(a) *The Theory* The essentials of the theory underlying the present test may be summarized as follows:

i A definition of a word consists of a fuzzy set of semantic features, of varying criteriality.
ii Semantic features can be represented (in some form of logical notation) as statements predicated of the definiendum.
iii Semantic features are not necessarily semantic primes, but may have quite a complex internal propositional structure of their own.
iv Semantic features include features specifying selection restrictions (e.g. the restriction that 'walk' is predicated of animate

beings with legs).

v Implicative relations may exist between semantic features (e.g. 'x is a vessel' implies 'x is a physical object').

vi The criteriality of semantic features can vary according to (a) context (b) idiolect.

(b) *Theory in relation to Testing, Description, and 'Real Meaning'*
In terms of the model outlined above, statements about meaning can be made at at least three different levels. On the testing level, an experimental result such as (V) may be said to be an approximative description of meaning in the form of a fuzzy set of features. Initially, the testing method itself must be evaluated for its conformity to intuitively arrived at statements of meaning. In improving the predictive value of the tests, various distorting factors inherent in the testing method must be investigated and allowed for. But even in an unimproved form, the testing method may be used for comparative statements, e.g. about dialectal variation in lexical meaning. In a more refined form, the method may hopefully be used with more confidence in absolute descriptions of meaning, where intuitive evidence gives no support.

Ideally, a lexical definition should be a description of a 'real meaning', which we may interpret, in psychological terms, as the knowledge by virtue of which a given native speaker is able to use and interpret the word in question in a manner which agrees (within certain margins of idiolectal variation) with a parallel knowledge possessed by other native speakers. For various reasons, this 'real meaning' is not precisely statable. One reason is that the number of features entering into the definition of a word (if we include any feature of above zero criteriality) is indefinitely large, and no criterion may be supplied for deciding when an exhaustive definition has been given. Another reason is that in theory, an index of criteriality may be given the value of any real number from 0 to 1 (assuming a conflation of positive and negative criteriality), so that observational criteriality is never likely to be more than an approximation to theoretical criteriality. A further reason is that any programme of experimentation which would permit a precise and valid numerical value to be assigned to criteriality seems beyond the realm of practical possibility.

Since it is not possible to give a satisfactory lexical definition on either the observational level or the theoretical level, an intermediate 'descriptive' level is called for. On this level, a lexical definition is given in a form which is claimed to be a significant approximation to the 'real meaning'. A descriptive definition can take a number of equally valid forms, but is obviously limited to the specification of an important subset of the features in the 'real meaning' definition (let us say those over a certain degree of criteriality) and to the specification of criteriality in an approximate form. I suggest, in fact, that a strict-categorial definition (as in componential analysis) is one justifiable form of approximate definition, whose theoretical power and simplicity may, in some circumstances, overrule considerations of accuracy. (A strict-categorial definition is a special case of a fuzzy-set definition, in which degree of membership is specified only in terms of the values 1 and 0). Since the linguist can only aspire to descriptive, approximative definitions, his object, in conducting semantic tests, is to verify descriptive criteriality against observational criteriality.

(c) *Sociolinguistic Variation* It is likely that idiolectal and sociolinguistic variation in the meaning of words tends to be inversely proportional to criteriality, i.e. that people tend to agree more about the 'core' than about the 'periphery' of a word's meaning. Therefore the more descriptive definition is limited to features of high criteriality, the more its validity is generalizable over linguistic populations and linguistic situations. The approximative nature of a descriptive definition, while weakening its claim to accuracy, strengthens its claim to generality. Hence an apparent objection to semantic tests -- that they merely average out the divergent responses of members of the informant sample, and so arrive at a 'consensus' result -- can be presented in a more favourable light. Certainly, the investigator must be alert to sociolinguistic variations within his sample population of speakers, and may well make the study of such variations a goal of his tests; but variations in scoring between one member of the sample and another are by no means necessarily attributable to idiolectal or sociolinguistic differences. A moderate dispersion of responses is to be expected in view of different informants' implementation of the scale (it is clear from the present test, for example, that some informants markedly favour extremes of the scale, and others favour the more neutral positions). But so long as such factors remain roughly constant in different samples, there is no reason to suspect the general reliability of the overall result. The representation of approximate criteriality as a real number i where $0 \leqslant i \leqslant 1$ (in accord with Zadeh's index of fuzzy-set membership) is by no means the only conceivable measurement of vagueness that might be used, and further consideration must be given to the finding of an appropriate mathematical model for representing relative degrees of criteriality. Strictly, the answer to the question posed in the title of this paper must be 'No'; yet if the goal of precise lexical definition has to be given up, the experimental techniques discussed here, designed to specify lexical meaning in terms of statistical approximation, provide a reasonable second best.

9. *Comparison of Labov's Method and the Linguistic-Stimulus Method*

The two methods of research into lexical semantics, Labov's visual-stimulus method and my linguistic-stimulus method, both suffer from the limitation of incompleteness, in that the criteria selected for study can only be a subset of those which might possibly play a significant part in the definition under investigation. Apart from this, the methods, as I have suggested, have complementary weaknesses and strengths, and should therefore be used conjunctively, rather than as alternatives. The basic difference between them is that Labov's technique is referential (denotative) in orientation, while the linguistic-stimulus technique is semantic (connotative) in orientation. That is, Labov's method (with slight modification) asks 'What values are placed on a set of criteria in deciding whether a given object can be named by a given word?' while my method (given a referential interpretation) asks 'What values are placed on a set of criteria in deciding whether a given word can name a given object?' the answers to these questions can differ, without being inconsistent. In Labov's tests, for example, the use of cups for holding flowers showed a stronger negative result than the use of cups for eating out of. This may be attributable to a greater referential overlap between *bowl* and *cup* than between *vase* and *cup*. Such considerations of alternative names are not

reflected in my platitude test, which shows the 'flowers' criterion and the 'food' criterion as having virtually the same negative value.

The greatest weakness of the visual-stimulus method, its virtual limitation to physical criteria, must be set against a complementary weakness of the linguistic-stimulus method in dealing with physical criteria which are not readily verbalizable (e.g. criteria for colour terms). A notable instance of this is the width-depth ratio criterion, which in the platitude test I attempted to evaluate by means of the three statements describing a cup as *deep/shallow/of middling depth*. Labov's 'ideal' 1:1 ratio for a cup would, I hoped, be corroborated by a high criteriality rating for statement 16 *A cup is a container (or vessel) of middling depth.* But although this statement received a higher rating than the other two, the test assigned it the low criteriality of .18. Thus while (as I have argued) Labov gives this criterion unjustified prominence, my test, mainly because of the statement's lack of clarity, appears to undervalue it. Both types of test show fairly obvious imperfections, but offer the promise that, with further refinement of experimental techniques, the 'measurement of meaning' can be placed on a sound empirical footing.

NOTES

1 See Labov (1973), also Anderson (1975), for a fruitful application of Labov's methods to the study of the acquisition of word meanings. An earlier study of semantic vagueness by Lehrer (1970) anticipates the quantification of features of meaning proposed in the present article.

2 To Wittgenstein (esp. 1953:66-80) and Quine (esp. 1960: Chs 2 and 4) belongs the credit for penetrating analyses of lexical vagueness and indeterminacy. In defence of linguists (such as the present writer) who have done their best to ignore vagueness for the sake of tidy analyses, it can be argued that not all lexical meanings are vague, and that indeed a substantial proportion of a language's lexicon may be successfully treated in strict-categorial terms.

3 E.g. in an article in the *TV Times* (August 23-9, 1975) (on Sir George Cayley's gliding experiments): 'Before then, the more adventurous had been airborne in balloons, but that was floating, not flying, *since balloons are lighter than air*'. The same article describes the glider as 'the first manned *aircraft*' and Cayley's reluctant collaborator, his coachman, as 'the world's first *pilot*' [my ital.] The writer here seizes on 'heavier than air' as criterial to his use of *fly*, *pilot* and *aircraft*, while ignoring the criterion of powered locomotion.

4 E.g. overheard in a pub: an argument on whether Muhammed Ali was 'boasting' in calling himself the greatest boxer; presumably the argument boiled down to whether 'self-praise' or 'unwarranted self-praise' was the correct definition of *boasting*.

5 See, for example, the many examples of semantic 'shift' in Waldron (1967 Ch. 7). A recent example is the way in which the

meaning of *assassination* seems to have shifted with the ascendancy of the expression *sectarian assassinations* with reference to murders in N. Ireland. Here the criterion that the assassinee should be an important personage is suppressed in favour of the criterion that the murder should have a political motive.

6 Labov's discussion of functional criteria (1973:359-60) is obscure. He seems to believe that functional criteria can in theory be quantified in the same sense as the width-depth ratio. But although one can envisage certain cases where referential vagueness of a physical character affects a functional criterion (e.g. does thick soup count as a liquid or solid food?), in general it is difficult to conceive of functional criteria in terms of a scale of referential values.

7 See Leech and Pepicello 1972:14-16.

8 The concept of fuzzy set-membership naturally implicates a parallel concept of fuzzy truth-values. Lakoff (1972:185) cites the examples:

a	A robin is a bird.	(true)
b	A chicken is a bird.	(less true than a)
c	A penguin is a bird.	(less true than b)
d	A bat is a bird.	(false, or at least very far from true)
e	A cow is a bird.	(false)

On fuzzy-set theory and fuzzy logic, see Zadeh 1972, Lakoff 1972, and the references cited in those articles.

9 The test informants were British 1st year undergraduates (who had not yet been exposed to semantics teaching). The statements were presented orally, at 15-second intervals, and in random order. The problem of polysemy was avoided by asking the students to concentrate on the 'normal kind of cup', ignoring bra-cups, egg-cups, football cups and the like. I am grateful to the informants for their cooperation, and to Caroline Clapham and R.D.McKinlay for advice on test design.

10 See Zadeh (1972:7).

11 The set-membership function would also predict a rather low frequency of response in the ±2 range, but this was not reflected in the scoring, no doubt chiefly because of a psychological tendency to avoid extremes of the scale, and therefore to select ±2 as a 'safer' alternative to ±3.

12 On the one-to-one correspondence between semantic features and statements, see Leech (1969:26-8, 74-5).

REFERENCES

Anderson, E.S. (1975) 'Cups and Glasses: Learning that Boundaries are Vague', *J.Child Language*, 2:79-103.

Labov, W. (1973) 'The Boundaries of Words and their Meanings', in C.-J.N.Bailey and R.W.Shuy eds, *New Ways of Analyzing Variation in English*. Washington D.C., Georgetown U.P.

Lakoff, G. (1972) 'Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts', in *Papers from the Eighth Regional Meeting*, Chicago Linguistic Society, 183-228.

Leech, G.N. (1969) *Towards a Semantic Description of English*. London, Longman.

(1970) 'On the Theory and Practice of Semantic Testing', *Lingua*, 24:343-64.

Leech, G.N. and Pepicello, W.J. (1972) 'Semantic Versus Factual Knowledge: An Experimental Approach', in D.Cohen ed., *Papers from the Symposium on Limiting the Domain of Linguistics*. Linguistics Department, University of Wisconsin-Milwaukee.

Lehrer, A. (1970) 'Indeterminacy in Semantic Description', *Glossa*, 4:37-109.

Quine, W.V. (1960) *Word and Object*. Cambridge (Mass), M.I.T. Press.

Waldron, R.A. (1967) *Sense and Sense Development*. London, Deutsch.

Wittgenstein, L. (1953) *Philosophical Investigations*. Oxford, Blackwell.

Zadeh, L.A. (1972) 'A Fuzzy-set-theoretic Interpretation of Linguistic Hedges', *J. of Cybernetics*, 2:4-34.